Best L_1-Approximation with Splines

by

Oleksandra Kharakuleva

Thesis
for the degree of
Master of Science

(Master i Anvendt matematikk og mekanikk)

Faculty of Mathematics and Natural Sciences
University of Oslo

November 2013

Det matematisk- naturvitenskapelige fakultet
Universitetet i Oslo

Acknowledgments

I would like to thank my supervisor Knut Mørken for his advices and patience, especially for finding the encouraging words, motivation and understanding.

I am thankfull to all my teachers through all these years at universities and schools. My special gratitude goes to my teacher Olga Blazhkiv for giving me a deep interest in math at school.

I am sincerely grateful to Helene, Goeun and Lena for being not just good friends for me, but also for all the help and feedback on this thesis. Thank you, Johnny, Yura and Natalia, for super support and being there for me. And of course, thank you, B1002, especially Zhiyan, Jarle, Andreas and Ida for all the laughs and a good mood during the last semester.

And my last but not least thank you goes to my dearest parents and my brother. I would not be here where I am now without your love, support and understanding. Thank you so much!

Oleksandra Kharakuleva
Oslo, November 2013

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The aim of this thesis is to develop an algorithm for solving the best approximation problem in the L_1 - norm with splines. Our study is based on the theory of L_1-approximation with polynomials as well as theory that explain how this can be extended to splines.

The reason why we focus on the L_1-norm is because it is a poorly developed subject compared to L_2 and L_∞, therefore, we think that it deserves more attention. This norm has many names in various fields, for example Manhattan norm, Sum of Absolute Difference, Mean- Absolute Error. The L_1-norm has a remarkable property that makes it highly suitable for fitting to discrete data that may have some large errors in data due to blunders. Namely, that the magnitude of a blunder makes no difference to the final approximation. The L_1 approximation is robust to unavoidable measurement errors.

Nowadays, the theory of approximation is applied more and more to resolve practical issues. Research shows that the most popular theories of approximations in the 80's were approximation by polynomials and rational approximation. However, at this time the theory of splines became a serious competitor to these approximations. And this is not because of a mere desire for new approaches, but due to many advantages compared to other theories. Among them are the following:

- local support of splines, i.e, the behavior of the spline in the neighborhood of a point does not effect the behavior of the spline in general, in contrast to polynomial interpolation;
- good approximation properties of splines, especially for non-smooth functions.

Splines possess many other noteworthy properties, some of which we will discuss below.

The theory of splines is a generalization of the theory of polynomials. In the first part of the thesis, we are going to study the theory of best L_1 approximation with polynomials. We will consider both the continuous and the discrete cases. Later we will also consider two practical algorithms for polynomial L_1-approximation, based on Lagrange interpolation and linear programming. The space of polynomials is an example of a so-called Chebyshev subspace, it is therefore necessary to discuss some of the main properties of such spaces that can be usefull for the spline case. We will introduce a more general class so-called weak Chebyshev spaces. Spline spaces are prototypes of weak Chebyshev spaces. We will give an algorithm for computing so-called canonical points for splines, based on the Newtons method for finding zeros of a system of equations.

Since we have used Chebyshev spaces, Chebyshev polynomials, which were named after Russian mathematician Chebyshev, it seems interesting to investigate russian names involved in the development of splines and approximation theory in general. We can name Korneychuk, Stechkin, Makarov, Zavjalov [14] for their contributions in spline theory, together with Goncharov [3], who worked mostly with approximation theories, including L_1-approximation. Most of the theory on best approximation and approximation in L_1-norm were taking from Powell [8], Nürnberger [6], Pinkus [7] and Cheney [1]. A number of valuable works on approximation theory and methods are written, for example the books of Watson, Strauss [11], Sommer [10]. Some of them, like Nürnberger [6], successfully combined the theory of L_1-approximation with spline theory. He worked with best approximation by functions from Chebyshev and spline spaces in both the uniform and L_1-norm. The theory on the relation between B-splines and Chebyshev sets was taken from [6]. Micchelli [4] also has many publications on approximations with polynomials and splines, he showed the existence and uniqueness of canonical points for Chebyshev spaces. The use of B-splines and its properties together with algorithm for computing B-spline was taken from Lyche and Mørken [12].

Chapter 2

Fundamental theory

As we stated before the main purpose of our work is to study L_1 approximation with splines space. And in order to do this, we will first study the theory of L_1 - approximation with polynomials and start by looking at the general interpolation problem, not just from a practical view, but also as well as an important theoretical tool.

2.1 Interpolation problem

Interpolation is the most common method of approximation of the functions. We consider Lagrange interpolation by spline and polynomial spaces.

Let $[a, b]$ be a closed and bounded interval. The space of continuous functions on $[a, b]$ is denoted by

$$\mathcal{C}[a, b] = \{f : [a, b] \to \mathbb{R} : f \text{continuous}\}.$$

The space of r-times continuously differentiable functions on $[a, b]$, where $r \in \mathbb{N}$ is denoted

$$\mathcal{C}^r[a, b] = \{f : [a, b] \to \mathbb{R} : f^{(r)} \in \mathcal{C}[a, b]\}.$$

We want to approximate a function $f \in [a, b]$ with a function g, such that g interpolates some certain points in the interval $[a, b]$. This leads to the following formulation of the interpolation problem.

Definition 2.1. *Let $\mathcal{G}$ be an n- dimensional subspace of $\mathcal{C}[a, b]$ and f be a function in $\mathcal{C}[a, b]$ and some points $t_1 < t_2 < \ldots < t_n$ of the interval $[a, b]$ be given. The Lagrange interpolation problem is to determine a function $g \in \mathcal{G}$ such that*

$$g(t_j) = f(t_j), \quad j = 1, \ldots, n. \tag{2.1}$$

If $\{g_1, \dots, g_n\}$ is a basis of $\mathcal{G}$ then we have to find coefficients $a_1, \dots, a_n$ such that

$$\sum_{i=1}^{n} a_i g_i(t_j) = f(t_j), \quad j = 1, \dots, n.$$

And this is equivalent to solving the following system of linear equations

$$\begin{pmatrix} g_1(t_1) & \dots & g_n(t_1) \\ \vdots & \ddots & \vdots \\ g_1(t_n) & \dots & g_n(t_n) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} f(t_1) \\ \vdots \\ f(t_n) \end{pmatrix}.$$

This can be written as $G\mathbf{a} = \mathbf{f}$ in vector form. This problem has a unique solution if and only if

$$\det(G) \neq 0,$$

where

$$\det(G) = \begin{vmatrix} g_1(t_1) & \dots & g_n(t_1) \\ \vdots & \ddots & \vdots \\ g_1(t_n) & \dots & g_n(t_n) \end{vmatrix}.$$

2.2 Chebyshev spaces

The so-called Chebyshev spaces play an important role in connection with this interpolation problems.

Definition 2.2. *An n-dimensional subspace $\mathcal{G}$ of $\mathcal{C}[a,b]$ is called a Chebyshev subspace or a Haar subspace if there exists a basis $g_1, \dots, g_n$ of $\mathcal{G}$ such that* $\det G > 0$*, for all $t_1 < \dots < t_n$ in $[a,b]$.*

One of the simplest examples on a Chebyshev subspace of $\mathcal{C}[a,b]$ is the subspace of polynomials of degree n,

$$\mathcal{P}_n = \{p : [a,b] \to \mathbb{R} : p(t) = \sum_{j=0}^{n} a_j t^j\}, \tag{2.2}$$

where $a_0, \dots, a_n \in \mathbb{R}$.

Theorem 2.3. *For an n-dimensional subspace $\mathcal{G}$ of $\mathcal{C}[a,b]$ the following statements are equivalent:*

- *For all functions $f \in \mathcal{C}[a,b]$ and all $t_1 < \dots < t_n$ in $[a,b]$ the Lagrange interpolation problem (2.1) has a unique solution from $\mathcal{G}$.*
- *$\mathcal{G}$ is a Chebyshev subspace.*

It also important to define what is the Chebyshev polynomials, and what properties they have.

Definition 2.4. *The function $T_n : [-1, 1] \to \mathbb{R}$, defined by*

$$T_n(t) = \cos(n \arccos t) \tag{2.3}$$

for all $t \in [-1, 1]$, is called the Chebyshev polynomial of degree n.

We can state the following properties of these polynomials.

Lemma 2.5. *The Chebyshev polynomials have the following properties:*

1. *The function T_n is a polynomial of degree n with $\|T_n\|_\infty = 1$.*
2. *$T_0(t) = 1, T_1(t) = t$ and*
$$T_{n+1} = 2tT_n(t) - T_{n-1}(t), \quad t \in [-1, 1], \quad n = 1, 2, \ldots. \tag{2.4}$$
3. *The zeros of T_{n+1} are*
$$t_i = \cos\left\{\frac{2(n+1-i)+1}{2(n+1)\pi}\right\}, \quad i = 1, \ldots, n+1. \tag{2.5}$$

We will state some other properties of these polynomials later.

In order to work with a spline spaces we have to introduce the more general class of Chebyshev spaces, namely a weak Chebyshev space. As we going to discuss later a weak Chebyshev space can be "approximated" by Chebyshev spaces.

Definition 2.6. *An n-dimensional subspace $\mathcal{G}$ of $\mathcal{C}[a, b]$ is called a weak Chebyshev subspace if there exists a basis $g_1, \ldots, g_n$ of $\mathcal{G}$ such that* $\det G \geq 0$*, for all $t_1 < \ldots < t_n$ in $[a, b]$.*

More details on the theory of approximation by Chebyshev and weak Chebyshev spaces can be found in books of Nürnberger [6], Powell [8], Cheney [1], Rice [9] and Micchelli [4].

2.3 Convexity

Convexity plays an important role in the approximation theory. We will use it later in order to discuss the unicity of the best approximation. Let a set of points $z_1, \ldots, z_m$ in $\mathbb{R}^n$ and a point z in $\mathbb{R}^n$ be given. Then the point z is a convex combination of the points $z_1, \ldots, z_m$, if

$$z = \sum_{j=1}^{m} t_j z_j, \tag{2.6}$$

where $t_j \geq 0$ for each j and $\sum_{j=1}^{m} t_j = 1$. It is called strict convex combination if none of the t_j vanishes [13]. For $m = 2$ the set of convex combinations of two points is just a line sengent that connects these points.

A subset V of $\mathbb{R}^n$ is called convex if, for every $x, y \in V$, the set V also contains all points on the line segment connecting x and y. In other words, $tx + (1-t)y \in V$, for every $t \in [0,1]$. As can be seen in the picture below, the set to the left is convex a set, but the one to the left is a non-convex set.

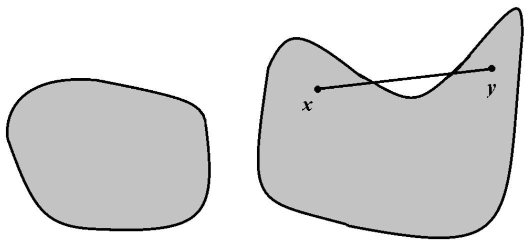

Figure 2.1: Example of a convex set (to the left) and a non-convex set (to the right).

2.4 Trapezoidal Rule

We will need a method for the numerical integration. The trapezoidal rule is based on linear interpolation of the function f in some points x_i, where $i \in \mathbb{N}$. The easiest example is to approximate the function $f \in [a,b]$ at the end points, i.e. $x_1 = a, x_2 = b$

$$\int_a^b f(x)dx \approx \frac{(b-a)}{2}\left(f(a) - f(b)\right). \tag{2.7}$$

The integral is equal to the area of a trapezoid with base $(b-a)$ times the average height $\frac{1}{2}\left(f(a) + f(b)\right)$. For better accuracy we subdivide the interval $[a,b]$ and assume that that $f_i = f(x_i)$ is known. Let $x_1 = a$, $x_n = b$, then $x_i = x_1 + ih$, where $h = \frac{b-a}{n-1}$ is a step length. The trapezoidal approximation for the i^{th} subinterval is

$$\int_{x_i}^{x_{i+1}} f(x)dx \approx \frac{h}{2}\left(f_i - f_{i+1}\right). \tag{2.8}$$

If we sum the contributions for each interval $[x_i, x_{i+1}]$, for $i = 1, \ldots, n$, then

$$\int_a^b f(x)dx \approx \frac{h}{2}\left(f(a) - f(b)\right) + \sum_{n=2}^{n-1} f_i. \tag{2.9}$$

For a more detailed description of this method and other methods of numerical integration one can check books of Mørken [5] and Dahlquit and Björck [2].

2.5 Summary

In this chapter we introduced the basic theory that is needed to understand the next chapters. We discussed one of the main tools of approximation, interpolation. We defined Chebyshev spaces and its main properties. Also we have introduced some main definitions of the convexity set and explained one of the methods for numerical integration.

The next two chapters include both theory on best approximation and some practical examples on L_1 approximation with interpolation.

Chapter 3

Approximation problem: Existence and uniqueness

In this chapter we will introduce some necessary theory on best approximation. We will discuss some basic properties of best approximation for both metric spaces and normed linear spaces and introduce some theorems on existence, unicity and continuity of the best approximation. We will also introduce the Haar conditions. Most of this theory behind this chapter is based on Powell [8].

3.1 Best approximation

First we will start with approximations in a metric space M. One of the main properties of a metric space is existence of a distance function $d(x, y)$, that is defined for all pairs of points (x, y) in M. It has the properties of reflexivity, positivity, symmetry and triangular inequality.

In most approximation problems there exists a suitable metric space that contains both f, where f is a data, function or more generally an element of a set, that is to be approximated and the set of approximations $\mathscr{A}$ [8]. We define $a^* \in \mathscr{A}$ to be a best approximation if the condition

$$d(a^*, f) \leq d(a, f) \tag{3.1}$$

holds for all $a \in \mathscr{A}$.

Theorem 3.1. *If $\mathscr{A}$ is a compact subset of a metric space M, then for every element f in M, there exists a best approximation a^* to f from $\mathscr{A}$.*

Best approximation may not exist if $\mathscr{A}$ is not compact. For example, if M is the Euclidean space $\mathbb{R}^2$, let $\mathscr{A}$ be the set of points that are strictly inside the unit circle. Then there are no best approximation to any point of M that is outside or on the unit circle [8].

The theory about metric space is not sufficiently enough for most of our work. That is why we choose to work with a normed linear space B. We recall from linear algebra that a linear space has the following three ingredients:

- a set E of elements (vectors and points);
- an operation in E called addition which obeys the usual rules of arithmetic;
- an operation of multiplying vectors by real numbers, obeying the usual rules of arithmetic.

A normed linear space is a linear space in which a real-valued function on vectors, called the norm, is defined. The norm is denoted by $\|\cdot\|$. It is known that every closed, bounded, finite-dimensional set in normed linear space is compact [1]. We assume that both f and $\mathscr{A}$ are contained in B.

The problem of best approximation in a normed linear space B can be also formulated as determination an element $a^* \in \mathscr{A}$ such that

$$\|f - a^*\| = \inf_{a \in \mathscr{A}} \|f - a\| = d(f, \mathscr{A}).$$

Theorem 3.2. *If $\mathscr{A}$ is a finite-dimensional linear space in a normed linear space B, then, for every $f \in B$, there exists an element of $\mathscr{A}$ that is a best approximation from $\mathscr{A}$ to f.*

We want to state that the best approximations from a finite-dimensional subspace of a strictly convex space are always unique. But first we will study some important theorems.

Theorem 3.3. *[8] Let $\mathscr{A}$ be a convex set of a normed linear space B, and let f be any point of B such that there exists a best approximation from $\mathscr{A}$ to f. Then the set of best approximation is convex.*

The uniqueness theorems require either $\mathscr{A}$ or the norm of the linear space B to be strictly convex. Both Powell [8] and Cheney [1] give theorems on convexity of approximation.

Theorem 3.4. *[8] Let $\mathscr{A}$ be a compact and strictly convex set of a normed linear space B. Then for all $f \in B$, there is just one best approximation from $\mathscr{A}$ to f.*

If there is a unique best approximation from $\mathscr{A}$ to f for all $f \in B$, then there must be a best approximation operator from B to $\mathscr{A}$ which must be a projection and we will denote it as X. We will show that very often this operator X must be continuous.

Theorem 3.5. *If $\mathscr{A}$ is a finite-dimensional linear space in a normed linear space B,such that for all functions in B, there is a unique best approximation from $\mathscr{A}$, $X(f)$ say. Then for all $f_1, f_2 \in B$*

$$|d(f_1, \mathscr{A}) - d(f_2, \mathscr{A})| \leq \|f_1 - f_2\|$$

and the operator X, defined by the best approximation condition, is continuous.

3.2 The Haar conditions

Powell [8] discusses the Haar conditions in order to introduce the characterization theorem for minimax approximation. We will refer to those properties that are relevant for L_1-approximation.

- If element of $\mathcal{P}_n$ has more than n-zeros, then it is identically zero.
- Let ξ_j, where $j = 1, 2, \ldots k$ be a set of distinct points in the open interval (a, b), where $k \leq n$. Then there exists an element of $\mathcal{P}_n$ that changes sign in these points, and has no other zeros. Moreover, there is a function in $\mathcal{P}_n$ that has no zeros in $[a, b]$.
- If a function in $\mathcal{P}_n$ that is not identically equal to zero, has j zeros, and if k of these zeros are interior points of $[a, b]$ at which the function does not change sign, then the number $j + k \leq n$.
- Let ξ_j, where $j = 0, 1, \ldots, n$ be any set of distinct points in $[a, b]$, and let ϕ_i, where $i = 0, 1, \ldots, n$ be any basis of $\mathcal{P}_n$. Then the $(n+1) \times (n+1)$ matrix whose elements have the values $\phi_i(\xi_j)$, where $i, j = 0, 1, 2, \ldots, n$ is non-singular.

An $(n + 1)$- dimensional linear subspace $\mathscr{A}$ of $\mathcal{C}[a, b]$ is said to satisfy the Haar condition if these four statements remain true when $\mathcal{P}_n$ is replaced by the set $\mathscr{A}$. Equivalently, any basis of $\mathscr{A}$ is called a Chebyshev set.

3.3 Summary

This chapter has given us some basic understanding of the existence and uniqueness of the best approximation in both metric and normed linear spaces. As we know, the three most frequently used norms are the L_p-norms where $p = 1, 2, \infty$. In the next chapter the least used norm L_1 of these three, will be studied more closely. If the linear space $\mathscr{A}$ satisfies the Haar condition, stated in this chapter, then theory on the best L_1-approximation from $\mathscr{A}$ to f can be explained and studied more.

Chapter 4

Best L_1 approximation with polynomials

Best approximation can be considered with respect to various norms. Often the choice of the norm depends on the given minimization problem. In this chapter we will take a closer look at approximations in the L_1-norm. We will state some important results on characterization and unicity of the best L_1-approximation with polynomials. We will discuss both the continuous and the discrete cases. Also we will introduce different methods and algorithms for approximation.

4.1 Continuous case

In most approximation problems the function f and the set of approximations $\mathscr{A}$ are in the space of continuous functions that are defined on the interval $[a, b]$, namely $\mathcal{C}[a, b]$. The L_1-norm in $\mathcal{C}[a, b]$ is defined by

$$\|f\|_1 = \int_a^b |f(x)|dx. \tag{4.1}$$

A best L_1 - approximation from a subset $\mathscr{A}$ of $\mathcal{C}[a, b]$ to a function f in $\mathcal{C}[a, b]$ is an element of $\mathscr{A}$ that minimizes the expression

$$\|f - p\|_1 = \int_a^b |f(x) - p(x)|dx, \quad p \in \mathscr{A}. \tag{4.2}$$

In order to introduce the characterization theorem, we consider first an example of approximation of a strictly monotone function f in $\mathcal{C}[a, b]$, by a constant function p, where the value of the constant is to be determined. The subset of approximations $\mathscr{A}$ is a linear space of dimension one. We will use the same example as Powell used in his book [8]. We want to find a value of a shaded area, that is the value of expression (4.2), see Figure. The approximation p is the function $p(x) = f(\xi)$, where f is a given function, $x \in [a,b]$ and $\xi - const$. We require the value of ξ that minimizes that area. One can see from the figure that, if we replace function $p(x) = f(\xi) + \epsilon$, where ϵ is very small, then we change the area of shaded regions. The area of the left-hand shaded region is approximately changed by $\epsilon(\xi - a)$, and the change of the right-hand shaded area is $-\epsilon(b - \xi)$, which gives a total change of about $2\epsilon(\xi - \frac{1}{2}[a+b])$. Therefore if $\xi < \frac{1}{2}[a+b]$, we can reduce $\|f - p\|_1$ by letting ϵ be positive, and, if $\xi > \frac{1}{2}[a+b]$, then there exists a negative value of ϵ that reduces the error. It follows that the required approximation is the constant function $p(x) = f(\frac{1}{2}(a+b))$, where $x \in [a,b]$. This approximation is optimal because the measures of the sets $\{x : f(x) < p(x)\}$ and $\{x : f(x) > p(x)\}$ are equal. This is an example of a condition for the best approximation that depends just on the sign of the error function. Another useful property of this example is that if we know in advance that f is monotone, then the calculation of $f(x)$ at a single point $x = \frac{1}{2}(a+b)$ provides all the data that are needed to determine the best approximation [8].

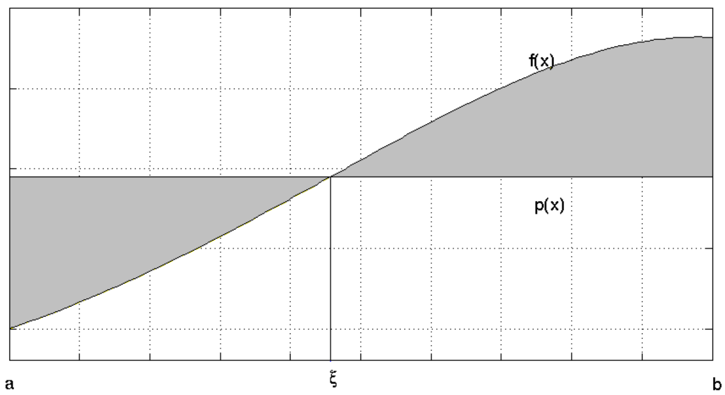

Figure 4.1: The value of $\|f - p\|_1$

Now we will introduce an extension of this example, a theorem, that gives the basic necessary and sufficient condition for a function p^* to be the best L_1 - approximation from $\mathscr{A}$ for a given f.

Theorem 4.1. *Let $\mathscr{A}$ be a linear subspace of $C[a,b]$. Let f be any function in $C[a,b]$, and let p^* be any element of $\mathscr{A}$, such that the set*

$$\mathcal{Z} = \{x : f(x) = p^*(x), a \le x \le b\} \tag{4.3}$$

is either empty or is composed of a finite number of intervals and discrete points. Then p^* is a best L_1 - approximation from $\mathscr{A}$ to f, if and only if the inequality

$$\left| \int_a^b s^*(x) p(x) dx \right| \le \int_{\mathcal{Z}} |p(x)| dx \tag{4.4}$$

is satisfied for all p in $\mathscr{A}$, where s^* is the sign function

$$s^*(x) = \begin{cases} -1, & f(x) < p^*(x) \\ 0, & f(x) = p^*(x) \qquad a \leq x \leq b. \\ 1, & f(x) < p^*(x), \end{cases} \tag{4.5}$$

This theorem can be used for calculation of the best approximation directly. If we consider the example above, the required approximation is the function $\{p^*(x) = f(\frac{1}{2}(a+b)); a \leq x \leq b\}$, because then the sign function (4.5) satisfies the characterization condition (4.4).

Theorem 4.2. *Let $\mathscr{A}$ be an $(n+1)$- dimensional linear subspace of $\mathcal{C}[a,b]$ that satisfies Haar condition, and let f be any function in $\mathcal{C}[a,b]$. If p^* is a best L_1 - approximation from $\mathscr{A}$ to f, and if the number of zeros of the error function*

$$e^*(x) = f(x) - p^*(x), \quad a \leq x \leq b, \tag{4.6}$$

is finite, then e^* changes sign at least $(n+1)$ times.

The next theorem tells us about the uniqueness of the L_1 - approximation, if the Haar condition is satisfied.

Theorem 4.3. *Let $\mathscr{A}$ be linear space of $\mathcal{C}[a,b]$ that satisfies the Haar condition. Then, for any function f in $\mathcal{C}[a,b]$ there is just one set L_1 - approximation from $\mathscr{A}$ to f.*

Most of algorithms for calculating the best L_1 - approximations aim to find the zeros of the error function. Often the number of zeros are exactly $(n+1)$, where $(n+1)$ is the dimension of the space $\mathscr{A}$. For example, this case occurs if $\mathscr{A}$ is the space $\mathcal{P}_n$, if f is in $\mathcal{C}[a,b]$, and if the derivative $f^{(n+1)}(x)$ is positive for all x in $[a,b]$. That is why the following theorem is useful.

Theorem 4.4. *Let $\mathscr{A}$ be an $(n+1)-$dimensional linear subspace of $\mathcal{C}[a,b]$ that satisfies the Haar condition, and let f be a function in $\mathcal{C}[a,b]$ such that the error function (4.6) has the exactly $(n+1)$ zeros, where p^* is the best L_1 - approximation from $\mathscr{A}$ to f. Then the position of the zeros does not depend on f.*

This theorem provides the main method for calculating the best L_1 - approximations to continuous functions. Assume that the error function changes sign only $(n+1)$ times. Since the zeros of the error function are independent of f, they may be found by detailed consideration of $\mathscr{A}$. We calculate the approximation by interpolation at these zeros, and then we check if its error function satisfies the assumption. If the assumption holds, then we got the required approximation. If not, then we need some other algorithm, for example a linear programming method, for more details see [13].

The next theorem gives us the necessary interpolation points in the special case when $\mathscr{A}$ is the space $\mathcal{P}_n$.

Theorem 4.5. *Let the conditions of Theorem 4.4 be satisfied, where $\mathscr{A}$ is the space $\mathcal{P}_n$ and where $[a, b]$ is the interval $[-1, 1]$. Then the zeros of the error function*

$$e^*(x) = f(x) - p^*(x), \quad 1 \leq x \leq -1, \tag{4.7}$$

have the values

$$\xi_i = \cos\left[\frac{(n+1-i)}{n+2}\right], \quad i = 0, 1, \cdots, n. \tag{4.8}$$

For a general interval $[a, b]$ the zeros of the error function are mapped to

$$\xi_i = \frac{1}{2}(a+b) + \frac{1}{2}(b-a)\cos\left[\frac{(n+1-i)}{n+2}\right], \quad i = 0, 1, \cdots, n. \tag{4.9}$$

Note that these points ξ_i are also known as the extreme points of the Chebyshev polynomial T_{n+2} on the interval $(-1, 1)$.

If we want to approximate the function f with the polynomial p of degree n. Then dimension to the space of approximations $\mathscr{A}$ is equal to $n+1$. This means that the error function $e^*(x)$ will have exactly $n+1$ zeros.

We want to check the last theorem with some examples. For this we define the set of points that have to be interpolated $\{x_i\}_{i=1}^{n+1}$ and let $f(x)$ be some function, such that $f(x) \in \mathcal{C}^{(n+1)}[-1, 1]$.

4.1.1 Examples

In the examples we are going to use the Bernstein basis, as a basis for polynomials. As we know, every polynomial in $\mathcal{P}_n[x_{min}, x_{max}]$ can be written as linear combination of Bernstein polynomials.

$$p(x) = \sum_{i=0}^{n} \widehat{B}_i^n(x) b_i,$$

where

$$\widehat{B}_i^n = B_i^n\left(\frac{x - x_{min}}{x_{max} - x_{min}}\right),$$

$$B_i^n(t) = \binom{n}{i} t^i (1-t)^{n-i}, \qquad i = 0, 1, \ldots, n.$$

To find b_i we have to solve the following system of equations

$$\begin{pmatrix} \widehat{B}_0^n(x_0) & \widehat{B}_1^n(x_0) & \cdots & \widehat{B}_n^n(x_0) \\ \widehat{B}_0^n(x_1) & \widehat{B}_1^n(x_1) & \cdots & \widehat{B}_n^n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{B}_0^n(x_n) & \widehat{B}_1^n(x_n) & \cdots & \widehat{B}_n^n(x_n) \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{pmatrix}.$$

Equidistant points	Chebyshev points	Points from Theorem 4.5
-1	-0.707106781186548	-0.500000000000000
1	0.707106781186547	0.500000000000000

Table 4.1: Table of the interpolation points for the approximation of the function $f(x) = x^2 + 1$ with the polynomials of degree $n = 1$.

As a test, we choose different sets of points: equidistant, Chebyshev points and points that give the best L_1 - approximation, i.e. the points $\{\xi_i, i = 0, 1, \dots, n\}$ given in Theorem 4.5. We want to check if the last set of points will minimize the expression $\|f - p\|_1 = \int_{-1}^{1} |f(x) - p(x)| dx$, for $p \in \mathscr{A}$.

We are going to make all figures and scripts for all our computations and analysis in an interactive environment as Matlab.

Example 4.6. *Let $f(x) = x^2 + 1$, for $x \in [-1, 1]$. We want to approximate this function by a polynomial $p \in \mathcal{P}_1$, i.e. p has degree $n = 1$. This means that our error function is going to have two zeros. This means that the polynomial $p(x)$ will interpolate two points. For the illustration see the Figure 4.2 and Table 4.1, the values of the minimized integrals can be found in Table 4.2*

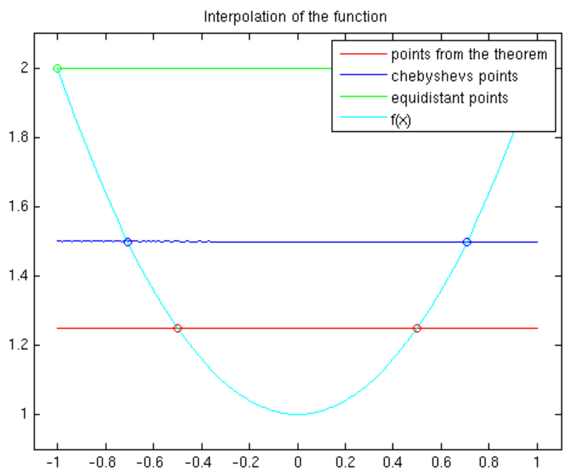

Figure 4.2: Interpolation of the function $f(x) = x^2 + 1$ on the interval $[-1, 1]$, with polynomials of degree $n = 1$. The points of interpolation are chosen to be equidistant, Chebyshev points, points from Theorem 4.5.

Equidistant points	Chebyshev points	Points from Theorem 4.5
1.333300000000001	6.095200e-01	5.000000e-01

Table 4.2: Table of the value of integrals $\|f-p\|_1 = \int_{-1}^{1} |f(x)-p(x)|dx$, where $p \in \mathscr{A}_2$

Equidistant points	Chebyshev points	Points from Theorem 4.5
0.367879441171442	0.363407118176516	0.356796028318389
0.655703035935713	0.655846360493777	0.659872590614888
0.969431932738860	0.971450457716832	0.980152211616769
2.718281828459046	2.711624962223608	2.701769152486852

Table 4.3: Coefficients b_i for polynomial $p(x) = \sum_{i=0}^{3} \widehat{B}_i^3(x) b_i$.

Example 4.7. *Let $f(x) = e^x$, for $x \in [-1,1]$. We want to approximate this function by a polynomial p, such that $p \in \mathcal{P}_3$, i.e. p has degree $n = 3$. This means that our error function is going to have four zeros, i.e. the polynomial $p(x)$ will interpolate four points. See Figure 4.3 for the illustration of the approximation. For the difference between coefficients of interpolation for different points see Table 4.3. The values of integrals $\|f-p\|_1 = \int_{-1}^{1} |f(x)-p(x)|dx$, where $p \in \mathscr{A}$ for different approximations are given in the Table 4.4.*

Equidistant points	Chebyshev points	Points from Theorem 4.5
0.009423185234593	6.862179e-03	5.430264e-03

Table 4.4: Table of the value of integrals $\|f-p\|_1 = \int_{-1}^{1} |f(x)-p(x)|dx$, where $p \in \mathscr{A}_4$.

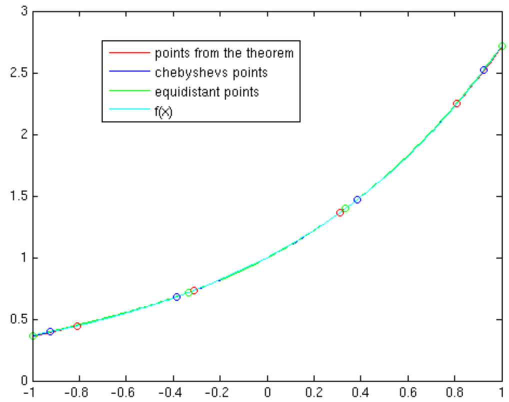

Figure 4.3: Interpolation of the function $f(x) = e^x$ on the interval $[-1, 1]$, with polynomials of degree $n = 3$. The points of interpolation are chosen to be equidistant, Chebyshev points, points from Theorem 4.5.

Equidistant points	Chebyshev points	Points from Theorem 4.5
-1.000000000000000	-0.965925826289068	-0.900968867902419
-0.600000000000000	-0.707106781186548	-0.623489801858733
-0.200000000000000	-0.258819045102521	-0.222520933956314
0.200000000000000	0.258819045102521	0.222520933956314
0.600000000000000	0.707106781186547	0.623489801858734
1.000000000000000	0.965925826289068	0.900968867902419

Table 4.5: Table of the interpolation points for the approximation of the function $f(x) = x^5 + 2$ with the polynomials of degree $n = 5$.

Example 4.8. *Let $f(x) = x^5 + 2$, for $x \in [-1, 1]$. We want to approximate this function by a polynomial p, such that $p \in \mathcal{P}_5$, i.e. p has degree $n = 5$. In other words, we will approximate the function $f(x)$ by the polynomial of the same degree $n = 5$. From the theory of approximation we know that in this case the error must be equal to zero. The interpolation points can be found in Table 4.6. On Figure 4.4 can be seen that all approximation curves coincide with the function and only interpolation points are visible. Table 4.6 gives the values of the integral $\|f - p\|_1 = \int\limits_{-1}^{1} |f(x) - p(x)|dx$, where function $f(x) = x^5 + 2$ and p is from the $\mathscr{A}_6$. The points from the theorem give best L_1-approximation, even though all integrals are equal to zero.*

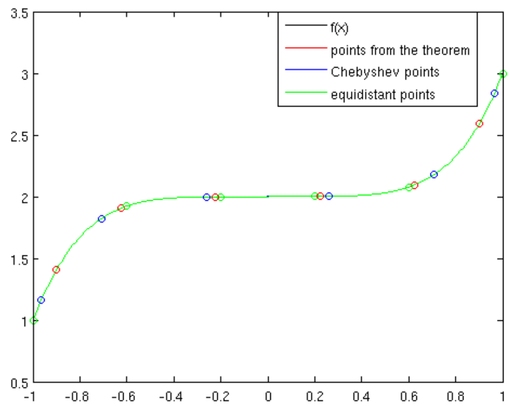

Figure 4.4: Interpolation of the function $f(x) = x^5 + 2$ on the interval $[-1, 1]$, with polynomials of degree $n = 5$. The points of interpolation are chosen to be equidistant, Chebysev points, points from Theorem 4.5.

Equidistant points	Chebyshev points	Points from Theorem 4.5
1.154632e-15	1.069145e-15	9.381385e-16

Table 4.6: Table of the value of integrals $\|f - p\|_1 = \int_{-1}^{1} |f(x) - p(x)|dx$, where $f(x) = x^5 + 2$ and $p \in \mathscr{A}_6$.

It is clear from all these examples that the points from Theorem 4.5 , or the canonical points, as we will call them later, give the best L^1 approximation. But as can be noticed from the last example, the error function is supposed to be zero, but we get an integral that is almost equal to zero. That is why it is natural to look on the condition number of the basis matrix B. The condition number of an $n \times n$ matrix B is $\mathrm{cond}(B) = \|B\|\|B^{-1}\|$. This number tells us how accurate we can expect the vector $\mathbf{b}$ when solving a system of equations $B\mathbf{b} = \mathbf{f}$. The numerical value of the condition number depends on the norm we use. The following table, Table 4.7 gives us the picture of how many digits of accuracy can be lost. We test all points and choose different norms. it can be seen that the Euclidian norm gives best accuracy.

	Condition number								
	Equidistant points			Chebyshev points			Canonical points		
$n \setminus p$	1	2	∞	1	2	∞	1	2	∞
1	1.0000e+00	1.0000e+00	1.0000e+00	1.4142e+00	1.4142e+00	1.4142e+00	2.0000e+00	2.0000e+00	2.0000e+00
2	2.5000e+00	2.3187e+00	3.0000e+00	3.3750e+00	2.7630e+00	3.6666e+00	5.0000e+00	4.0000e+00	5.0000e+00
3	6.0000e+00	5.0541e+00	5.6666e+00	6.6101e+00	5.3116e+00	6.3524e+00	9.6175e+00	7.8576e+00	9.3333e+00
4	1.7515e+01	1.1802e+01	1.5222e+01	1.4127e+01	1.0380e+01	1.4477e+01	2.1418e+01	1.5453e+01	2.1000e+01
5	4.2775e+01	2.8642e+01	3.3533e+01	2.8097e+01	2.0473e+01	2.6454e+01	4.2263e+01	3.0522e+01	3.9600e+01
6	1.2151e+02	7.1146e+01	8.9240e+01	5.8216e+01	4.0600e+01	5.7696e+01	9.0023e+01	6.0557e+01	8.5800e+01
7	3.0190+02	1.7945e+02	2.1023e+02	1.1607e+02	8.0771e+01	1.0802e+02	1.7894e+02	1.2060e+02	1.6342e+02
8	8.4470e+02	4.5752e+02	5.5976e+02	2.3728e+02	1.6099e+02	2.3046e+02	3.7325e+02	2.4086e+02	3.4728e+02
9	2.1259e+03	1.1757e+03	1.3673e+03	4.7355e+02	3.2127e+02	4.3754e+02	7.4434e+02	4.8196e+02	6.6650e+02
10	5.8966e+03	3.0401e+03	3.6480e+03	9.6128e+02	6.4156e+02	9.2131e+02	1.5345e+03	9.6553e+02	1.3996e+03
20	1.0366e+08	4.8212e+07	5.5004e+07	1.0131e+06	6.5385e+05	9.4303e+05	1.6985e+06	1.0141e+06	1.4566e+06
30	1.9472e+12	8.7462e+11	9.7746e+11	1.0487e+09	6.6876e+08	9.6581e+08	1.8066e+09	1.0560e+09	1.5002e+09
40	4.1545e+16	1.4413e+16	2.0265e+16	1.0799e+12	6.8443e+11	9.8914e+11	1.8936e+12	1.0932e+12	1.5407e+12
50	1.6294e+18	1.8123e+17	6.9178e+17	1.10959e+15	7.0136e+14	1.0129e+15	1.9835e+15	1.1271e+15	1.5911e+15
60	1.5744e+18	4.7345e+17	8.3327e+17	5.8599e+17	6.6257e+17	5.3576e+17	3.7129e+17	2.2491e+17	2.9189e+17
70	6.6950e+18	5.3646e+17	1.8851e+18	3.2191e+19	1.0732e+18	1.7402e+19	1.0970e+19	1.0240e+18	5.0183e+18
80	1.0301e+19	3.7712e+18	8.0227e+18	3.0458e+21	1.0255e+18	9.7395e+20	4.5186e+20	7.3124e+17	1.0766e+20
90	3.2064e+19	1.0601e+18	2.4521e+19	7.2152e+23	9.4887e+18	1.1296e+23	2.3104e+23	1.1450e+19	4.1795e+22
100	1.9990e+19	1.6146e+18	1.7864e+19	2.0331e+24	7.7293e+19	5.0073e+23	5.9191e+24	3.1145e+18	7.9641e+23

Table 4.7: Table of the condition numbers in norms $p = 1, 2, \infty$ for the equidistant points, Chebyshev points and canonical points for different degrees n.

4.2 Discrete case

To find the L_1 - approximation for a discrete form of the minimization problem we require the function from the space $\mathscr{A}$ that minimizes the expression

$$\sum_{t=1}^{m} w_i \left| f(x_t) - p(x_t) \right|, \quad p \in \mathscr{A}, \tag{4.10}$$

where $\{x_t; t = 1, 2, \cdots, m\}$ is the set of data points in $[a, b]$ and w_i - some fixed positive weight[1]. There is a characterization theorem for the discrete case as well.

Theorem 4.9. *Let the function values $\{f(x_t); t = 1, 2, \ldots, m\}$ and fixed positive weights $\{w_t; t = 1, 2, \ldots, m\}$ be given. Let $\mathscr{A}$ be a linear space of functions that are defined on the point set $\{x_t; t = 1, 2, \cdots, m\}$. Let p^* be any element of $\mathscr{A}$, let $\mathcal{Z}$ contain the points of $\{x_t; t = 1, 2, \cdots, m\}$ that satisfy the condition*

$$p^*(x_t) = f(x_t), \tag{4.11}$$

and let s^* be a sign function

$$s^*(x) = \begin{cases} -1, & f(x_t) < p^*(x_t) \\ 0, & f(x_t) = p^*(x_t) \\ 1, & f(x_t) < p^*(x_t), \end{cases} \qquad t = 1, 2, \cdots m. \tag{4.12}$$

Then p^* is a function in $\mathscr{A}$ that minimizes the expression

$$\sum_{t=1}^{m} w_t \left| f(x_t) - p(x_t) \right|, \quad p \in \mathscr{A}, \tag{4.13}$$

if and only if the inequality

$$\sum_{t=1}^{m} w_t s^*(x_t) p(x_t) \leq \sum_{x_t \in \mathcal{Z}} w_t \left| p(x_t) \right| \tag{4.14}$$

holds for all p in $\mathscr{A}$.

There are different methods of getting the best discrete L_1 - approximation. One of the most popular and most used methods of calculation the approximation is a linear programming problem. One of the main difference of the linear programming problem and the interpolation method for continuous case is that the zeros of the error function are unknown.

[1] we define $w_i = 1$

So our aim now is to compare if both methods will give us the same results. We can check it by comparing the coefficient of the polynomial. The following linear programming problem computes coefficients to the polynomial interpolation. Let us discuss this method in more detail. We let $\{\phi_i; i = 0, 1, \cdots n\}$ be a basis of the space $\mathscr{A}$ of approximations. The expression (4.13) that needs to be minimized , can be written in the form

$$\sum_{t=1}^{m} w_t \left| f(x_t) - \sum_{i=0}^{n} \lambda_i \phi_i(x_t) \right|, \tag{4.15}$$

where the parameters λ_i, where $i = 0, 1, \dots n$ are some variables of the linear programming calculation. We need also to introduce two new variables for each data point, which we will call $\{u_t; t = 1, 2, \dots m\}$ and $\{v_t; t = 1, 2, \dots m\}$. The linear optimization problem gets the following form,

$$\begin{aligned} \min \quad & \textstyle\sum_{t=1}^{m} w_t(u_t + v_t) \\ \text{s.t.} \quad & f(x_t) - \textstyle\sum_{i=0}^{n} \lambda_i \phi_i(x_t) \leq u_t, \\ & f(x_t) - \textstyle\sum_{i=0}^{n} \lambda_i \phi_i(x_t) \geqslant v_t, \qquad t = 1, 2, \dots m. \\ & u_t \geqslant 0, \\ & v_t \geqslant 0. \end{aligned} \tag{4.16}$$

Here we have $(2m + n + 1)$ linearly independent constrains and equally many variables.

Our main problem of minimization is to find coefficients $\{\lambda_i; i = 0, 1, \dots, n\}$. For the discrete case, we will use both methods of finding those coefficients and will compare results. Our aim is to see if both methods give the same approximation. As before we use Matlab for the interpolation method and we will use Cplex for the optimization problem. We will use the same functions as in the previous examples.

Example 4.10. *Let $f(x) = x^2 + 1$, for $x \in [-1, 1]$. The following Table 4.8 gives the value of coefficients of interpolation for different points.*

Equidistant points	Chebyshev points	Points from the Theorem 4.5	points from Cplex
2	1.5	1.25	1.25
2	1.5	1.25	1.25

Table 4.8: Table of coefficients for different points, $f(x) = x^2 + 1$

Here we can see that we got the same coefficients in both methods.

Example 4.11. *Let $f(x) = e^x$, for $x \in [-1, 1]$. To compare coefficients, see Table 4.9.*

Equidistant points	Chebyshev points	Points from the Theorem 4.5	points from Cplex
0.367879441171442	0.363407118176516	0.356796028318389	0.35685
0.655703035935713	0.655846360493777	0.659872590614888	0.65984
0.969431932738860	0.971450457716832	0.980152211616769	0.98008
2.718281828459046	2.711624962223608	2.701769152486852	2.7019

Table 4.9: Table of the coefficients of interpolation for different points, $f(x) = e^x$

We can try an example were we can be sure that optimal solution will be zero. We know that if we interpolate a function of degree n by a polynomial of the same degree or higher, then the optimal solution will agree with the function and the error will be equal to zero.

Example 4.12. *Let $f(x) = x^5 + 2$, for $x \in [-1, 1]$. We want to approximate this function by a polynomial of the degree $n = 7$. We get the following coefficients, see Table 4.10.*

Points from Theorem 4.5	points from Cplex
0.999999999999998	1
2.428571428571434	2.4286
1.952380952380935	1.9524
1.857142857142889	1.8571
2.142857142857109	2.1429
2.047619047619065	2.0476
1.571428571428568	1.5714
3.000000000000000	3

Table 4.10: Table of the coefficients of interpolation for different points, $f(x) = x^5 + 2$

The solution has error equals to zero and the minimized integral is equal to 1.078582e-15. If we compare our coefficients, we can see that they are almost equal. That is because of the round-off errors. It looks like Matlab and Cplex use different tolerance.

4.3 Summary

In this chapter we discussed best L_1-approximation for both the continuous case and the discrete case. First we introduced the theory on the best L_1-approximation in the continuous case. We stated a theorem about the uniqueness of the best L_1-approximation. The most interesting theorem is about the position of the zeros of the error function which do not depend on f. From this we obtained a method of L_1 approximation that is based on Lagrange interpolation. We implemented this method in Matlab, and got good results. We chose the Bernstein basis, that is a spline basis in a polynomial case. Our results confirmed the theory. Thereafter we introduce some basic theory in the discrete case. An algorithm of minimization was proposed. Even though it is given for a discrete case, we extended it to the continuous case. The linear programming problem was more difficult to implement, compare to interpolation problem. However, the results were the same with the exception that commercial linear programming solver does not allow very small tolerances.

In the next chapter we will look more closely on B-splines and their properties. We will discuss the space of splines and see that they are prototypes of a weak Chebyshev space, just like space of polynomials $\mathcal{P}_n$ is a prototype of a Chebyshev space.

Chapter 5

Basic properties of splines and B-splines

In this chapter we will consider best approximation by Chebyshev spaces in the L_1- norm. We will also give some theorems that concern weak Chebyshev spaces. We will introduce B-splines and discuss their main properties. We will also discuss matrix representation of splines and the Schoenberg and Whitney theorem.

5.1 B - splines. Basic properties

B-splines made their first appearance in Schoenberg's 1946 paper on the equidistant data by analytic functions. There is no doubt that B-splines appeared earlier in literature. They play a prominent role already in Favard's work, and Schoenberg has always claimed that they were already known to Laplace. But it is in Schoenberg's paper that they were thought important enough to be given a name, "basic k^{th}- order spline curves". This was the same paper in which Schoenberg introduced splines.

Let p be a nonnegtive integer, and let $\mathbf{t}$ be a nondecreasing sequence of real numbers of length at least $p+2$, called the knot vector. The jth B-spline of degree p with knots t_j is defined by

$$B_{j,p,\mathbf{t}}(x) = \frac{x - t_j}{t_{j+p} - t_j} B_{j,p-1,\mathbf{t}}(x) + \frac{t_{j+1} - x}{t_{j+1+p} - t_{j+1}} B_{j+1,p-1,\mathbf{t}}(x), \tag{5.1}$$

for all real numbers x, with

$$B_{j,0,\mathbf{t}}(x) = \begin{cases} 1, & \text{if } t_j \le x \le t_{j+1}, \\ 0, & \text{otherwise.} \end{cases} \tag{5.2}$$

The B-spline $B_{j,p}$ depends only on the knots $(t_k)_{k=j}^{j+p+1}$. This knot interval $[t_j, t_{j+p+1}]$ is called support. We can say that B-splines have "minimal" support. So if the knot vector is given by $\mathbf{t} = (t_j)_{j=1}^{n+p+1}$ for some positive integer n, we can form n B-splines $\{B_{j,p}\}_{j=1}^{n}$ of degree p associated with this knot vector. A linear combination of B-splines, or a spline function, is a combination of B-splines on the form

$$f = \sum_{j=1}^{n} c_j B_{j,p}, \tag{5.3}$$

where $\mathbf{c} = (c_j)_{j=1}^{n}$ are n real numbers. Let $\mathbf{t} = (t_j)_{j=1}^{n+p+1}$ be a nondecreasing sequence of real numbers. The linear space of all linear combinations of these B-splines is the spline space $S_{p,\mathbf{t}}$ defined by

$$S_{p,\mathbf{t}} = \operatorname{span}\{B_{1,p}, \dots, B_{n,p}\} = \{\sum_{j=1}^{n} c_j B_{j,p} | c_j \in \mathbb{R} \text{ for } 1 \leq j \leq n\}. \tag{5.4}$$

An element $f = \sum_{j=1}^{n} c_j B_{j,p}$ of $S_{p,\mathbf{t}}$ is called a spline function of degree p with knots $\mathbf{t}$ and $(c_j)_{j=1}^{n}$ are B-spline coefficients of f. The following lemma give us some basic properties of splines.

Lemma 5.1. *Let $\boldsymbol{t} = (t_j)_{j=1}^{n+p+1}$ be a knot vector for splines o degree p with $n > p+1$, and let $f = \sum_{j=1}^{n} c_j B_{j,p}$ be a spline in $S_{p,\boldsymbol{t}}$. Then f has following properties:*

1. *If x lies in the interval $[t_\mu, t_{\mu+1}]$ for some μ in range $p+1 < \mu < n$then*

$$f(x) = \sum_{j=\mu-p}^{\mu} c_j B_{j,p}(x). \tag{5.5}$$

2. *If $z = t_{j+1} = \cdots = t_{j+p} < t_{j+p+1}$ for some j in the range $1 \leq j \leq n$ then $f(z) = c_j$.*

3. *If z occurs m times in $\boldsymbol{t}$ then f has continuous derivatives of order $0, \dots, p-m$ at z.*

The next theorem gives us the matrix representation of B-splines.

Theorem 5.2. *Let $\boldsymbol{t} = (t_j)_{j=1}^{n+p+1}$ be a knot vector for B-splines of degree p, and let μ be an integer such that $t_\mu < t_{\mu+1}$ and $p+1 \leq \mu \leq n$. For each positive integer k with $k \leq p$ we define the matrix $\boldsymbol{R}_k^\mu(x) = \boldsymbol{R}_k(x)$ by*

$$\boldsymbol{R}_k(x) = \begin{pmatrix} \frac{t_{\mu+1}-x}{t_{\mu+1}-t_{\mu+1-k}} & \frac{x-t_{\mu+1-k}}{t_{\mu+1}-t_{\mu+1-k}} & 0 & \cdots & 0 \\ 0 & \frac{t_{\mu+2}-x}{t_{\mu+2}-t_{\mu+2-k}} & \frac{x-t_{\mu+2-k}}{t_{\mu+2}-t_{\mu+2-k}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{t_{\mu+k}-x}{t_{\mu+k}-t_\mu} & \frac{x-t_\mu}{t_{\mu+k}-t_\mu} \end{pmatrix} \tag{5.6}$$

Then for x in the interval $[t_\mu, t_{\mu+1})$, the $p+1$ B-splines $\{B_{j,p}\}_{j=1}^{\mu}$ of degree p that are nonzero on this interval can be written

$$\boldsymbol{B}_p = \begin{pmatrix} B_{\mu-p,p} & B_{\mu-p+1,p} & \dots & B_{\mu,p} \end{pmatrix} = \boldsymbol{R}_1(x)\boldsymbol{R}_2(x)\dots\boldsymbol{R}_p(x). \tag{5.7}$$

If $f = \sum_j c_j B_{j,p}$ is a spline in $S_{p,\boldsymbol{t}}$ and x is restricted to the interval $[t_\mu, t_{\mu+1})$, then $f(x)$ is given by

$$f(x) = \boldsymbol{R}_1(x)\boldsymbol{R}_2(x)\dots\boldsymbol{R}_p(x)c_0, \tag{5.8}$$

where c_0 is given by $c_0 = \begin{pmatrix} c_{\mu-p,p} & c_{\mu-p+1,p} & \dots & c_{\mu,p} \end{pmatrix}$. The matrix $\boldsymbol{R}_k(x)$ is called a B-spline matrix.

This theorem shows how one polynomial piece of a spline can be built up, by multiplying and adding together (via matrix multiplications) certain linear polynomials. This representation is only an alternative way to write the recurrence relation (5.1) [12]. The advantage of such matrix representation is that all the recursive steps are captured in one equation, that makes it very convenient. Tom Lyche og Knut Mørken in their book [12] give two algorithms for evaluating a spline. We can accumulate the matrix products from left to right or from right to left. We use the following algorithm for evaluating a spline.

Algorithm 5.3. *Let the polynomial degree p, the knots $t_{\mu-p+1} \leq t_\mu \leq t_{\mu+1} \leq t_{\mu+p}$ and a number x in $[t_\mu, t_{\mu+1})$ be given and set $\boldsymbol{B}_0 = 1$. After evaluation of the products*

$$\boldsymbol{B}_k(x)^T = \boldsymbol{B}_{k-1}(x)^T\boldsymbol{R}_k(x), \quad k = 1, \dots, p, \tag{5.9}$$

the vector $\boldsymbol{B}_p(x)$ will contain the value of the $p+1$ B-splines at x,

$$\boldsymbol{B}_p(x) = \begin{pmatrix} B_{\mu-p,p}(x) & B_{\mu-p+1,p}(x) & \dots & B_{\mu,p}(x) \end{pmatrix}^T \tag{5.10}$$

There is also a general spline interpolation problem that can be defined as following.

Problem 5.4. *Let there be given data $(x_i, f_i)_{i=1}^m$ and a spline space $S_{p,\boldsymbol{t}}$ whose knot vector $\boldsymbol{t} = (t_i)_{i=1}^{m+p+1}$ satisfies $t_{i+p+1} > t_i$, for $i = 1, \dots, m$. Find a spline g_f in $S_{p,\boldsymbol{t}}$ such that*

$$g_f(x_i) = \sum_{j=1}^{m} c_j B_{j,p}(x_i) = f_i \quad i = 1, \dots, m. \tag{5.11}$$

The next Theorem 5.5 gives necessary and sufficient conditions for this system to have a unique solution.

Theorem 5.5. *The matrix $(B_{j,p}(x_i))_{i,j=1}^m$ is nonsingular if and only if its diagonal is positive, i.e.,*

$$B_{j,p}(x_i) > 0, \quad i = 1, \dots, m. \tag{5.12}$$

5.2 Splines and Weak Chebyshev spaces

The theory of Chebyshev spaces is not applicable to spline spaces. However, spline spaces are prototypes of the more general class of weak Chebyshev spaces. A fundamental result says that weak Chebyshev spaces can be "approximated" by Chebyshev spaces. Therefore, although there are some crucial differences between Chebyshev and weak Chebyshev spaces, certain properties can be carried over to weak Chebyshev spaces. In particular, for every continuous function, there exists a best uniform approximation from a given weak Chebyshev space such that the error has a classical alternation property [6]. As we have noticed before, the existence of best approximation is guaranteed for every finite-dimensional space, but best approximations do not need to be unique in general. However, it is known that best L_1-approximations from Chebyshev spaces are always unique.

Theorem 5.6. *Let $\mathcal{G}$ be a Chebyshev space of $\mathcal{C}[a,b]$. Then for every function $f \in \mathcal{C}[a,b]$, there exists a unique best L_1- approximation from $\mathcal{G}$.*

This result was showed by Jackson in 1930 [6]. It is interesting to notice that in contrast to uniform approximation, the converse of the Theorem5.6 is not true. For example, best L_1 approximations from a spline space are always unique (see Theorem 5.12 in 3.2), but spline spaces are not Chebyshev spaces, because, as we said earlier, spline spaces are prototypes of weak Chebyshev spaces.

First of all, let us discuss polynomial splines. Let points $a = t_0 < t_1 \ldots < t_n < t_{n+1} = b$ and an integer $p \geq 1$ be given. We call

$$S_{p,\boldsymbol{t}} = S_p(t_1, \ldots, t_n) = \{s \in \mathcal{C}^{p-1}[a,b] : s|_{[t_i,t_{i+1}]} \in \mathcal{P}_p, i = 0, \ldots n\} \quad (5.13)$$

the space of polynomial splines of degree p with n fixed knots $t_1, \ldots, t_n$.

There is a basis of a given spline space consisting of polynomials and truncated power functions, such that the dimension of $S_{p,\boldsymbol{t}}$ is $n + p + 1$. This is important, because the next theorem, due to Schoenberg and Whitney, says that spline spaces are weak Chebyshev spaces.

Theorem 5.7. *The space $S_p(t_1, \ldots, t_n)$ is an $(n + p + 1)$- dimensional weak Chebyshev subspace of $\mathcal{C}[a,b]$.*

The proof of this theorem is based on the well-known Rolle's Theorem and facts about the dimension of the basis formed by truncated power functions.

An interesting observation is made, that for functions from the so called convexity cone, best L_1- approximations from a given Chebyshev space can be computed by solving a Lagrange interpolation problem. The points which correspond to such interpolation problems are called canonical points.

Definition 5.8. *Let $\mathcal{G}$ be an n- dimensional subspace of $\mathcal{C}[a,b]$. We call points $t_1 < \ldots < t_r$ in (a,b), where $1 \leq r \leq n$, canonical points for $\mathcal{G}$ if*

$$\sum_{i=0}^{r} (-1)^i \int_{t_i}^{t_{i+1}} g(t)dt = 0 \tag{5.14}$$

for all $g \in \mathcal{G}$, where $t_0 = a$ and $t_{r+1} = b$.

In 1977 Micchelli [4] showed the existence and uniqueness of canonical points with the following theorem.

Theorem 5.9. *[6] For every n- dimensional Chebyshev subspace of $\mathcal{C}[a,b]$, there exists a unique set of n canonical points.*

It is interesting to notice that Hobby and Rise showed that for every arbitrary n- dimensional subspace of $\mathcal{C}[a,b]$ there exists a set of r canonical points, where $1 \leq r \leq n$.

It is relevant for us to investigate the relationship between best L_1 approximation and interpolation at the canonical points. The convexity cone plays a central role in this context.

Definition 5.10. *Let $\mathcal{G}$ be a Chebyshev subspace of $\mathcal{C}[a,b]$. The set*

$$C(\mathcal{G}) = \{f \in \mathcal{C}[a,b] : span(\mathcal{G} \cup f) \text{is a Chebyshev subset of } \mathcal{C}[a,b]\}. \tag{5.15}$$

is called convexity cone of $\mathcal{G}$.

This means that the function $f \in \mathcal{C}^{n+1}[a,b]$ with $f^{n+1}(t) \neq 0$ for all $t \in (a,b)$ belongs to convexity cone $C(\mathcal{P}_n)$. The following theorem, due to Micchelli [4], shows that for a function from a convexity cone, its best L_1-approximations from a given Chebyshev space can be computed by interpolation at canonical points. A similar result was proved by Bernstein for spaces of polynomials in 1926.

Theorem 5.11. *Let $\mathcal{G}$ be an n- dimensional Chebyshev subspace of $\mathcal{C}[a,b]$ and $t_1 < \ldots < t_n$ in (a,b) be the canonical points for $\mathcal{G}$. Then for every function $f \in C(\mathcal{G})$ its best L_1- approximation g_f from $\mathcal{G}$ is uniquely determined by*

$$g_f(t_i) = f(t_i), \quad i = 1, \ldots, n. \tag{5.16}$$

By the result of Bernstein, the canonical points of spaces of polynomials are explicitly known. See Theorem 4.5 from the chapter about polynomial interpolation.

5.3 Best L_1 - Approximation by Weak Chebyshev Spaces

We noted before that the best L_1 approximations from Chebyshev spaces are always unique. This result does not hold for the more general class of weak Chebyshev spaces. But the following theorem due to Strauss and Galkin shows that global unicity of best L_1 approximations hold for spline spaces.

Theorem 5.12. *For every function in $\mathcal{C}[a,b]$ there exists a unique L_1 - approximation from $S_p(t_1, \dots t_n)$*

Now we will introduce some of Micchelli's results [4] on existence of canonical points for weak Chebyshev spaces.

Theorem 5.13. *For every n-dimensional weak Chebyshev subspace of $\mathcal{C}[a,b]$, there exists a set of n canonical points.*

We already stated about the uniqueness of determination canonical points for a given Chebyshev space in Theorem 5.9. Micchelli developed sufficient conditions for the uniqueness and poisedness[1] of canonical points. Sommer gave the following version.

Theorem 5.14. *Let $\mathcal{G}$ be an n-dimensional weak Chebyshev subspace $\mathcal{C}[a,b]$ such that for every function in $\mathcal{C}[a,b]$, there exists a unique best L_1 approximation from $\mathcal{G}$. Then there exists a unique set of n canonical points of $\mathcal{G}$ which is poised with respect to $\mathcal{G}$.*

Corollary 5.15. *[6] For the space $S_{p,\mathbf{t}}$ there exists a unique set of $p+n+1$ canonical points which is poised with respect to $S_{p,\mathbf{t}}$.*

Miccheli proved some special relationship between best L_1- approximation and interpolation. He showed that best L_1-approximation for functions from the convexity cone of weak Chebyshev spaces can be obtained by Lagrange interpolation.

If $\mathcal{G}$ is a weak Chebyshev subspace of $\mathcal{C}[a,b]$ then the set

$$K(\mathcal{G}) = \{f \in \mathcal{C}[a,b] : \text{span}(\mathcal{G} \cup \{f\})\text{is a weak Chebyshev space of } \mathcal{C}[a,b]\}$$

is called the convexity cone of $\mathcal{G}$.

Theorem 5.16. *Let $\mathcal{G}$ be an n-dimensional weak Chebyshev subspace of $\mathcal{C}[a,b]$. If the set $\{t_1, \dots, t_n\}$ of canonical points of $\mathcal{G}$ is poised with respect to $\mathcal{G}$, then every function $f \in K(\mathcal{G})$ has a unique best L_1-approximation g_f from $\mathcal{G}$ an g_f is uniquely determined by*

$$g_f(t_i) = f(x_i), \quad i = 1, \dots, n. \tag{5.17}$$

[1] A subset $t_1, \dots, t_n$ of $[a,b]$ is called poised with respect to n-dimensional subspace $\mathcal{G}$ if $det(g_i(t_j)) \neq 0$, where $g_1, \dots, g_n$ is a basis for $\mathcal{G}$

Micchelli combined Theorem 5.12, Theorem 5.14 and Theorem 5.16 to the following corollary.

Theorem 5.17. *For every function $K(S_{p,t})$, the best L_1-approximation from $S_{p,t}$ is uniquely determined by Lagrange interpolation in canonical points.*

5.4 Summary

In this chapter we studied B-splines and their properties. We stated the algorithm for computing spline function, that we need for the implementation part. We studied properties of a weak Chebyshev space. We stated that L_1-approximations for functions from the convexity cone can be obtained by Lagrange interpolation. In the next chapter we will suggest the algorithm for computing canonical points, i.e. zeros of an error function.

Chapter 6

Computing canonical points

In this chapter we are going to present an algorithm for computing the canonical points. We are going to define the basic problem and the method a solution.

6.1 Defining the problem

First of all, let us define the problem that we want to solve. As we know from the theory of polynomials and its best L_1 approximation, due to Theorem 4.4, if $\mathcal{G}$ is an $(n+1)-$dimensional linear subspace of $\mathcal{C}[-1,1]$ that satisfies the Haar condition, the error function (4.6) has exactly $(n+1)$ zeros. If f is a function in $\mathcal{C}$ and p^* is the best L_1 approximation from $\mathcal{G}$ to f, then the positions of the zeros do not depend on f. These zeros of the error function $e^*(x) = f(x) - p^*(x)$, for $1 \leqslant x \leqslant -1$, have the values

$$\xi_i = \cos\left[\frac{(n+1-i)}{n+2}\right], \quad i = 0, 1, \cdots, n, \tag{6.1}$$

see Theorem 4.5. As we mentioned before these points are the canonical points. By Definition 5.2, if $\mathcal{G}$ is an n- dimensional subspace of $\mathcal{C}[-1,1]$, then the canonical points $t_1 < \ldots < t_r$ for $\mathcal{G}$ in $(-1,1)$, where $1 \leq r \leq n$, are the solution of the equation

$$\sum_{i=0}^{r}(-1)^i \int_{t_i}^{t_{i+1}} g(t)dt = 0 \tag{6.2}$$

for all $g \in \mathcal{G}$, where $t_0 = -1$ and $t_{r+1} = 1$. In this chapter we want to compute the canonical points sufficient to for a basis of $\mathcal{G}$.

6.1.1 Equations

We have a vector of unknowns, and this means that we are going to get a system of a non-linear equations. We are going to derive a numerical method for finding a solution. The solution of equation (6.2) can be found by Newton's method. We choose this method, because it converges faster than other iteration methods [5].

Newton's method solves the system of non-linear equations, which amounts to find the zeroes of a continuously differentiable function $\mathbf{G} : \mathbb{R}^n \to \mathbb{R}^n$. We start with some start vector $\mathbf{t}_0 = (t_{00}, t_{01}, \ldots, t_{0n})$, then we compute the zero vector $\mathbf{t}_1 = (t_{10}, t_{11}, \ldots . t_{1n})$ of a tangent at $\mathbf{t}_0$. We are going to use the following formula:

$$\mathbf{t}_i = \mathbf{t}_{i-1} - (\nabla J_G(\mathbf{t}_{i-1}))^{(-1)}\mathbf{G}(\mathbf{t}_{i-1}), \tag{6.3}$$

where $\nabla J_G(\mathbf{t}_{i-1})$ is the Jacobian matrix, and $\mathbf{G}(\mathbf{t}) \in \mathcal{G}$.

The Jacobian matrix $\nabla J_G(\mathbf{t})$ is the matrix of all first-order partial derivatives of a vector-valued function. For some $\mathbf{G}(\mathbf{t}) = (G_1(\mathbf{t}), G_2(\mathbf{t}), \ldots, G_{n+1}(\mathbf{t}))$, where $\mathbf{t} = (t_j)_{j=1}^{n+1}$, the Jacobian matrix can be written in the following form

$$\nabla J_G(\mathbf{t}) = \begin{pmatrix} \frac{\partial G_1}{\partial t_1} & \frac{\partial G_1}{\partial t_2} & \cdots & \frac{\partial G_1}{\partial t_{n+1}} \\ \frac{\partial G_2}{\partial t_1} & \frac{\partial G_2}{\partial t_2} & \cdots & \frac{\partial G_2}{\partial t_{n+1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial G_{n+1}}{\partial t_1} & \frac{\partial G_{n+1}}{\partial t_2} & \cdots & \frac{\partial G_{n+1}}{\partial t_{n+1}} \end{pmatrix} \tag{6.4}$$

What do we need to calculate canonical points of the spline? First of all, we need to know the degree p of the interpolating spline. Secondly, we have to define how many points we want to interpolate, i.e. number of canonical points n. The following formula will help us to determine the function $\mathbf{G}(\mathbf{t})$.

$$\mathbf{G}(\mathbf{t}) = \int_{-1}^{t_1} g(t)dt - \int_{t_1}^{t_2} g(t)dt + \ldots + (-1)^{n-1}\int_{t_n}^{1} g(t)dt, \tag{6.5}$$

where $g(t)$ is a spline function

$$g(t_i) = \sum_{j=1}^{n} c_j B_{j,p}(t_i) \tag{6.6}$$

with spline coefficients (c_i) and knot vector $(\tau_i)_{i=1}^{n+p+1}$. We are going to use numerical integration to find these definite integrals. We will use the trapezoidal rule, the same rule as in the polynomial case.

The disadvantage of the Newton's method is the calculation of the partial derivatives. But in our case, this works fine, because of the possibility to determine the partial derivatives explicit. One can easily notice from equations (6.5), that the partial derivatives of the function $\mathbf{G}$ are equal to

$$\begin{aligned} \frac{\partial G}{\partial t_1} &= 2g(t_1), \\ \frac{\partial G}{\partial t_2} &= -2g(t_2), \\ \dots &\quad \dots \\ \frac{\partial G}{\partial t_n} &= (-1)^{n-1}2g(t_n). \end{aligned} \tag{6.7}$$

This means that we can find Jacobian matrix explicit. To make our life even easier we choose the coefficients of the B-splines in such way, that function $\mathbf{G}(t) = (G_1, \dots, G_n)(t)$ can be defined in the following way

$$\begin{aligned} G_1 &= \int_{-1}^{t_1} B_{1n}(s)ds - \int_{t_1}^{t_2} B_{1n}(s)ds + \ldots + (-1)^{n-1} \int_{t_n}^{1} B_{1n}(s)ds \\ G_2 &= \int_{-1}^{t_1} B_{2n}(s)ds - \int_{t_1}^{t_2} B_{2n}(s)ds + \ldots + (-1)^{n-1} \int_{t_n}^{1} B_{2n}(s)ds \\ &\vdots \qquad \ddots \quad \ddots \\ G_i &= \int_{-1}^{t_1} B_{in}(s)ds - \int_{t_1}^{t_2} B_{in}(s)ds + \ldots + (-1)^{n-1} \int_{t_n}^{1} B_{in}(s)ds \\ &\vdots \qquad \ddots \quad \ddots \\ G_n &= \int_{-1}^{t_1} B_{nn}(s)ds - \int_{t_1}^{t_2} B_{nn}(s)ds + \ldots + (-1)^{n-1} \int_{t_n}^{1} B_{nn}(s)ds. \end{aligned} \tag{6.8}$$

We define function $G_i(\mathbf{t})$ to be function $\mathbf{G}(\mathbf{t})$ with spline coefficients equal to zero except the $c_i = 1$. This representation saves us much time and makes it easier to find the Jacobian matrix. If we look closer on the way the Jacobian matrix can be presented in our case, we can see that it is transposed to the matrix that represents the B-spline basis. But there is a difference - each column is multiplied by $(-1)^{j-1} * 2$.

6.1.2 Algorithm

So what is the algorithm for finding such canonical points for splines?

One of important phases of making the algorithm is to define what variables do we need for finding the solution. In our case, the most important variable is the start vector $\mathbf{t_0}$, our first guess of the solution. The length of the vector gives us the number of canonical points n. Also we need to know the degree p of the spline, so that we define the knot vector $\boldsymbol{\tau}$ that is $p+1$-regular knot[1].

In this algorithm J_{ij} is the Jacobian matrix, $\{B_j\}_t$ B-spline basis and $B_i(s)$ is the B-spline evaluated at some point s. To compute Jacobian matrix we

[1] A $p+1$ extended knot for which $t_1 = t_{p+1}$ and $t_{n+1} = t_{n+p+1}$.

Algorithm 1 Calculate the canonical points

Require: spline space $S_{p,\tau}$; initial value $\mathbf{t_0}$;
number of points $n = dim(S_{p,\tau})$;
tolerance $\epsilon = 10^{-15}$;
$\mathbf{t} = \mathbf{t_0}$;
$stopNum = 1$;
while $stopNum > \epsilon$ **do**
The i^{th} component of the $\mathbf{G}(t_1, \dots, t_n)$:
$G_i(t_1, \dots, t_n) = \int_{-1}^{t_1} B_i(s)ds - \int_{t_1}^{t_2} B_i(s)ds + \dots + (-1)^{n-1} \int_{t_n}^{1} B_i(s)ds$;
$J_{ij} = (-1)^{j-1} * 2B_{ji}$;
$\mathbf{tt} = \mathbf{t} - (J)^{-1} * G(\mathbf{t})$;
$stopNum = norm(\mathbf{tt} - \mathbf{t})/norm(\mathbf{tt})$;
$\mathbf{t} = \mathbf{tt}$;
end while

use

$$G_i(t_1, \dots, t_n) = \int_{-1}^{t_1} B_i(s)ds - \int_{t_1}^{t_2} B_i(s)ds + \dots + (-1)^{n-1} \int_{t_n}^{1} B_i(s)ds.$$

This looks easy, but it requires some work, since we have to compute a B-spline by the algorithm from the Chapter 5 about splines. Then then each integral is computed by numerical integration, using the trapezoidal rule, changing the bounds of integration. We stop our algorithm and get the result when

$$\frac{\|\mathbf{tt} - \mathbf{t}\|}{\|\mathbf{tt}\|} \leq 10^{-15}. \tag{6.9}$$

where $\mathbf{tt}$ is the latest approximation, while $\mathbf{t}$ is the previous one.

6.2 Summary

In this chapter we introduced the algorithm for finding the canonical points for splines. For this we used Newton's method for finding the zeros of the system of equations. The advantage of this method is its convergence and the explicity of the Jacobian matrix. We tried briefly to describe the algorithm that can be implemented for different software. In the next chapter we will present some results and observations that was made computing canonical points by this method.

Chapter 7

Experiments and results

In this chapter we are going to show and analyze some our results, that in our opinion, are important and can be useful for further research.

Our aim is to find the canonical points of splines. We are going to use the algorithm described in the previous chapter. As we stated before all our experiments we are going to be conducted in Matlab. We will just briefly state some necessary and important results.

7.1 Polynomial case

Our first experiment is to compute canonical points for a spline tin the pure polynomial case. That means that we have to define a spline space $S_{p,\tau}$, τ is a $p+1$- regular knot-vector with no interior knots. As we decided before we are going to work on the interval $[-1, 1]$. We decided to illustrate examples with five canonical points. For this we have to determine our start vector $\mathbf{t}_0$ and after some iterations get the sequence that converges to the roots of the system of equations. We just choose some random points on the interval $[-1, 1]$. Working with splines the Whitney -Schoenberg theorem plays one of the important roles. This case is a polynomial case, so if everything is correct we should get the canonical points for the polynomials, see Theorem 4.5. First let us calculate the canonical points for polynomials.

Canonical points
-0.866025403784439
-0.500000000000000
0.000000000000000
0.500000000000000
0.866025403784439

Table 7.1: Canonical points from the Theorem 4.5 for approximation space $\mathscr{A}_5$.

We choose degree $p = 4$, some start vector $\mathbf{t}_0 = [-0.9, -0.5, -0.1, 0.4, 0.8]$, and 5 -regular vector $\tau = [-1, -1, -1, -1, -1, 1, 1, 1, 1, 1]$. We use our Algorithm 6.1, we get the solution of our system of equation that is equal to the canonical points in polynomial case. We made some experiments with the different values of the start vector $\mathbf{t}_0$. We can illustrate our results in the following Table 7.2.

Start vector	Number of iterations
$[-1, -0.5, 0, 0.5, 1]$	8
$[-1, -0.6, -0.1, 0.5, 1]$	8
$[-0.7, -0.5, 0.1, 0.5, 0.6]$	11
$[-0.9, -0.4, -0.1, 0.5, 0.8]$	7
$[-0.8, -0.5, 0, 0.5, 0.8]$	7

Table 7.2: Table of the different start vectors and the number of iterations needed to get the solution given in Table 7.1 with tolerance 10^{-15}.

As we can see different start vectors give us the same result. This shows good convergence of Newtons method. We are using the tolerance 10^{-15} to stop the iterations.

7.2 One inner knot

We decided to check what points we get if we choose the degree of the spline equal to $p = 3$, with the same start vector $\mathbf{t}_0$. Then we have to adjust our knot vector to have length of $n + p + 1 = 9$, and be 4 regular,i,e. knot vector is going to have one inner knot. The most natural way to choose the inner knot at the point 0, that is in the middle of our interval. But we want also to investigate the cases when the inner knot is not in the middle. The results of these experiments can be seen in Table 7.3, Table 7.4, Table 7.5.

Start vector	$\tau_5 = -0.6$	Number iterations
$[-1, -0.5, 0, 0.5, 1]$	$[-0.9366, -0.6823, -0.2044, 0.3664, 0.8251]$	8
$[-0.7, -0.5, 0.1, 0.5, 0.6]$	singular matrix	
$[-1, -0.6, -0.1, 0.5, 1]$	$[-0.9366, -0.6823, -0.2044, 0.3664, 0.8251$	8
$[-0.9, -0.4, -0.1, 0.5, 0.8]$	$[-0.9366, -0.6823, -0.2044, 0.3664, 0.8251$	15
$[-0.8, -0.5, 0, 0.5, 0.8]$	singular matrix	

Table 7.3: Table of the different start vectors and canonical vectors for splines with a knot vector $\tau = [-1, -1, -1, -1, -0.6, 1, 1, 1, 1]$.

Start vector	$\tau_5 = 0$	Number iterations
$[-1, -0.5, 0, 0.5, 1]$	$[-0.8660, -0.5000, -0.0000, 0.5000, 0.8660]$	8
$[-1, -0.6, -0.1, 0.5, 1]$	$[-0.8660, -0.5000, -0.0000, 0.5000, 0.8660]$	8
$[-0.7, -0.5, 0.1, 0.5, 0.6]$	$[-0.8660, -0.5000, -0.0000, 0.5000, 0.8660]$	11
$[-0.9, -0.4, -0.1, 0.5, 0.8]$	$[-0.8660, -0.5000, -0.0000, 0.5000, 0.8660]$	8
$[-0.8, -0.5, 0, 0.5, 0.8]$	$[-0.8660, -0.5000, -0.0000, 0.5000, 0.8660]$	7

Table 7.4: Table of the different start vectors and canonical vectors for splines with a knot vector $\tau = [-1, -1, -1, -1, 0, 1, 1, 1, 1]$.

Start vector	$\tau_5 = 0.5$	Number iterations
$[-1, -0.5, 0, 0.5, 1]$	$[-0.8305, -0.3852, 0.1723, 0.6490, 0.9221]$	8
$[-1, -0.6, -0.1, 0.5, 1]$	$[-0.8305, -0.3852, 0.1723, 0.6490, 0.9221]$	8
$[-0.7, -0.5, 0.1, 0.5, 0.6]$	singular matrix	
$[-0.9, -0.4, -0.1, 0.5, 0.8]$	singular matrix	
$[-0.8, -0.5, 0, 0.5, 0.8]$	$[-0.8305, -0.3852, 0.1723, 0.6490, 0.9221]$	10

Table 7.5: Table of the different start vectors and canonical vectors for splines with a knot vector $\tau = [-1, -1, -1, -1, 0.5, 1, 1, 1, 1]$.

These observations again show good convergence of Newton's method. But the most interesting a value observation is the case when we have one inner knot in the middle of interval. The experiments show that we get the same canonical points as in the polynomial case of one degree higher.

7.3 Two inner knots

This observation force us to go futher and check if we increase number of inner knots to two, but decrease the degree of the spline to $p = 2$. We will check two cases, the first if we take a knot vector with the inner knots of multiplicity two, i.e $\tau = [-1, -1, -1, 0, 0, 1, 1, 1]$; an the second case if we take two uniformly distributed points on the interval $[-1, 1]$, as inner knots of our knot vector, i.e $\tau = [-1, -1, -1, -1/3, 1/3, 1, 1, 1]$. We will test same start vectors as in the previous experiments. See Table 7.6, and Table 7.7

Start vector	Canonical points	Number iterations
$[-1,-0.5,0,0.5,1]$	$[-0.8683,-0.5040,0.0000,0.5040,0.8683]$	8
$[-1,-0.6,-0.1,0.5,1]$	$[-0.8683,-0.5040,0.0000,0.5040,0.8683]$	8
$[-0.7,-0.5,0.1,0.5,0.6]$	$[-0.8683,-0.5040,0.0000,0.5040,0.8683]$	10
$[-0.9,-0.4,-0.1,0.5,0.8]$	$[-0.8683,-0.5040,0.0000,0.5040,0.8683]$	8
$[-0.8,-0.5,0,0.5,0.8]$	$[-0.8683,-0.5040,0.0000,0.5040,0.8683]$	7

Table 7.6: Table of the different start vectors and canonical vectors for splines with a knot vector $\tau = [-1,-1,-1,0,0,1,1,1]$

Start vector	Canonical points	Number iterations
$[-1,-0.5,0,0.5,1]$	$[-0.8406,-0.4545,-0.0000,0.4545,0.8406]$	8
$[-1,-0.6,-0.1,0.5,1]$	$[-0.8406,-0.4545,-0.0000,0.4545,0.8406]$	12
$[-0.7,-0.5,0.1,0.5,0.6]$	$[-0.8406,-0.4545,-0.0000,0.4545,0.8406]$	13
$[-0.9,-0.4,-0.1,0.5,0.8]$	$[-0.8406,-0.4545,-0.0000,0.4545,0.8406]$	12
$[-0.8,-0.5,0,0.5,0.8]$	$[-0.8406,-0.4545,-0.0000,0.4545,0.8406]$	7

Table 7.7: Table of the different start vectors and canonical vectors for splines with a knot vector $\tau = [-1,-1,-1,-1/3,1/3,1,1,1]$

As we can see we got some canonical points, that are quite close to the canonical points that we want to get, but not close enough to see any connection to polynomial case.

7.4 Discussion

Let us make some observations.

- no inner knots. Then our method works perfectly well. We get the same values for the canonical points as in the polynomial case. We noticed that result does not depend on the start vector, because the solution always converged.

- one inner knot. If the inserted knot is 0, then we get some interesting observation. The experiments show that we get the same canonical points as in the polynomial case of one degree higher.

  If the inserted knot is not in the middle, we get some different values. It is difficult to choose the start vector because in some cases we get a singular matrix. In our case one of the rows of the Jacobian matrix became zero.

- two inner knots. We tried two uniformly spaces points from the end-points,as well as point 0 of multiplicity two. We got some canonical points, but we can not find a simple link to polynomial case.

degree	Equidistant points	Chebyshev points	Points from Theorem 4.5
1	7.791674e-02	3.963239e-02	4.970348e-02
2	7.791674e-02	4.937191e-02	3.963239e-02
3	9.853882e-04	6.129598e-04	6.371599e-04
4	7.911993e-04	6.345586e-04	5.028020e-04
5	6.283470e-06	3.774660e-06	3.835403e-06
6	4.686227e-06	3.825985e-06	3.020590e-06
7	2.678263e-08	1.326907e-08	1.340416e-08
8	1.905628e-08	1.338617e-08	1.053962e-08
9	8.312094e-11	3.043534e-11	3.063301e-11
10	5.783911e-11	3.046935e-11	2.414286e-11
10	5.783911e-11	3.046935e-11	2.414286e-11
20	3.206426e-14	8.007619e-16	7.897813e-16
30	2.713370e-11	1.078688e-15	1.230157e-15
40	1.216320e-08	1.391028e-15	1.588512e-15
50	2.644679e-08	1.946067e-15	1.781933e-15
60	6.169630e-08	2.852868e-15	2.516344e-15
70	3.554488e-07	3.659855e-14	1.240970e-14
80	7.952090e-07	1.168651e-11	9.031547e-14
90	1.022251e-06	4.914889e-10	3.332768e-10
100	1.366243e-06	7.065178e-10	2.870551e-09

Table 7.8: Table of the value of minimized integrals between function from non-convexity cone $f = \sin x$, where $x \in [-1, 1]$ and its approximation.

It seems that it is important to mention one more algorithm for solving best L_1-approximation problem with splines. It can be solved as linear programming problem. In stead of Bernstein basis, that we used before to illustrate some examples, we can use a spline basis and solve the linear optimization problem (4.2). We tested this algorithm for polynomial case (no inner knots) and got the same solution as for interpolation problem.

One more test was made. We stated before that this theory works just for functions from the convexity cone. On the example we wanted to check if by interpolating canonical points for polynomials, see the Theorem 4.5 we will get the best approximation in L_1-norm. The value of the minimized integral can be seen in Table 7.8. The canonical points for polynomials do not minimize the integral best. We leave this question for further research.

Chapter 8

Summary

This thesis presents the study of the theory on the best L_1-approximation with splines. Our intention was to extend the theory of approximation with polynomials to splines. As a result, we developed an algorithm for solving the best approximation problem in L_1-norm with splines. The algorithm is based on the Newton's method. A number of interesting observations were made during the experiments. The first step of the experiment was to test whether the same canonical points for the spline in the pure polynomial case could be achieved. Further, we decided to decrease the degree of a spline and to adjust a knot vector by inserting a inner knot at the point zero. The resulting canonical points were the same as in the polynomial case of one degree higher. The next step was to test whether we can get the same canonical points by decreasing the degree of a spline and increasing the number of inner knots. However, our experiments showed that this holds just for one inner knot.

The work can be extended with some theoretical proofs to this observation. It would be interesting to investigate if the canonical points for splines are extreme points of Chebyshev splines, just like in the polynomial case canonical points are extreme points for Chebyshev polynomials T_{n+2}. Also our work could be extended to the theory of best L_1-approximations with spline surfaces.

In addition, for the further work, we could investigate how our results can be used to minimize L_0-norm, that is the solution of compressed sensing problem.

Bibliography

[1] E. W. Cheney. *Itroduction to approximation theory*. McGraw-Hill, New York, 1966.

[2] Å. Björck G. Dahlquist. *Numerical Methods in Scientific Computing*. Society for Industrial and Applied Mathematics, Philadelphia, 2008.

[3] V. L. Goncharov. *Theory of interpolation and approximation of functions*. State publisher of technical and theoretical literature, Moskva, 1954.

[4] C. A. Micchello. Best l^1 approximation by weak chebyshev systems and the uniqueness of interpolating perfect splines. *Journal of Approximation Theory*, 19:1–14, 1977.

[5] Knut Mørken. *Numerical Algorithms and Digital Representation*. Centre of Mathematics for Applications, University of Oslo, Oslo, 2013.

[6] G. Nürnberger. *Approximation by spline functions*. Springer-Verlag, Berlin, Heidelberg, New York, 1989.

[7] A. Pinkus. *On L^1 approximation*. Cambridge University Press, New York, NY, 1989.

[8] M. J. D. Powell. *Approximation theory and methods*. Cambridge University Press, New York, NY, 1996.

[9] J. D. Rice. *The approximation of functions*. Addison-Wesley publishing company, Reading, Massachusetts, 1964.

[10] M. Sommer. Weak chebyshev spaces and best l_1-approximation. *Journal of Approximation Theory*, 39:54–71, 1983.

[11] H. Strauss. Best l_1-approximation. *Journal of Approximation Theory*, 41:297–308, 1984.

[12] K. Mørken T. Lyche. *Spline methods*. Centre of Mathematics for Applications, University of Oslo, Oslo, 2011.

[13] R. J. Vanderbei. *Linear programming Foundations and Extensions Third Edition*. Springer, Princeton University, USA, 2010.

[14] V. A. Skorospelov Y. S. Zavjalov, V. A. Leus. *Methods of spline functions*. Nauka, Moskva, 1980.